Resolving the true band gap of ZrNiSn half-Heusler thermoelectric materials

Jennifer Schmitt,†a Zachary M. Gibbs,†b G. Jeffrey Snyder,*c and Claudia Felser*d

Cite this: Mater. Horiz., 2015, 2, 68

N-type XNiSn (X = Ti, Zr, Hf) half-Heusler (HH) compounds possess excellent thermoelectric properties, which are believed to be attributed to their relatively high mobility. However, p-type XNiSn HH compounds have poor figures of merit, $zT$, compared to XCoSb compounds. This can be traced to the suppression of the magnitude of the thermopower at high temperatures. $E_g = 2S_{\text{max}} T_{\text{max}}$ relates the band gap to the thermopower peak. However, from this formula, one would conclude that the band gap of p-type XNiSn solid solutions is only one-third that of n-type XNiSn, which effectively prevents p-type XNiSn HHs from being useful thermoelectric materials. The study of p-type HH Zr$_x$Sc$_{1-x}$NiSn solid solutions show that the large mobility difference between electrons and holes in XNiSn results in a significant correction to the Goldsmid–Sharp formula. This finding explains the difference in the thermopower band gap between n-type and p-type HH. The high electron-to-hole weighted mobility ratio leads to an effective suppression of the bipolar effect in the thermoelectric transport properties which is essential for high $zT$ values in n-type XNiSn ($X = Ti, Zr, Hf$) HH compounds.

Introduction

Thermoelectric (TE) materials convert thermal gradients to electric potential gradients, and thus they hold promise for generating electricity from waste heat in automobiles. The maximum efficiency of these materials is determined by their dimensionless figure of merit, $zT = S^2T/\kappa\rho$, where $S$ is the thermopower (absolute value of the Seebeck coefficient), $\rho$ is the electrical resistivity, and $\kappa$ is the total thermal conductivity, which is the sum of the lattice part $\kappa_l$ and the electronic part $\kappa_e$. Thus, good TE materials possess high thermopower, low electrical resistivity, and low thermal conductivity. Half-Heusler (HH) compounds with the general formula XNiSn ($X = Ti, Zr, Hf$) have generated significant interest as a promising class of materials for thermoelectric applications because of their high thermopower. The numerous possibilities to manipulate each of the three lattice sites provide an excellent opportunity to influence the electronic and thermal transport properties. Through isoelectronic alloying on the $X$ site or partial substitution on the Ni site, the thermal conductivity can be reduced owing to lattice strain and mass fluctuations by point-defect scattering of phonons. A phase separation in various HH solid solutions results in remarkably low thermal conductivities and enhanced electrical properties.

Much of the existing literature focuses on the thermoelectric characterization of n-type XNiSn and p-type XCoSb solid solutions, whereas studies on p-type XNiSn-based HHs are far fewer in number. The most promising p-type HHs are found in the XCoSb system ($X = Ti, Zr, Hf$). For the construction of thermoelectric modules, it is desirable for the n-type and p-type...
materials to have similar chemical, thermal, and mechanical properties. In order to be most compatible with high-performance HH n-type materials that are currently being explored for modules,\textsuperscript{21} a suitable p-type analog in the XNiSn system is necessary. Some work has already been done to this end. For example, multiple authors have studied the substitution of Ni by Co and demonstrated that a conversion from n- to p-type behavior for the XNiSn system could be achieved.\textsuperscript{22,23} Likewise, \textit{Horyn et al}.\textsuperscript{24} studied the effects of substituting Sc for Ti or Zr up to 350 K; the resulting compounds showed promising positive thermopower at room temperature.

The electronic transport properties of semiconducting materials, such as thermopower and electrical conductivity, are reflections of the electronic energy band structure. A crucial parameter that determines the thermoelectric performance is the band gap, \(E_g\). For narrow band-gap semiconductors, the onset of bipolar conduction (both electrons and holes contributing) occurs at lower temperatures, and their compensating thermopower lead to a lower temperature peak in the overall thermopower and therefore low \(zT\) values.

In this work, we aim to extend the existing studies of p-type XNiSn HH compounds by thoroughly investigating Sc substitutions in Zr\(_{1-x}\)Sc\(_x\)NiSn. Through Sc substitution, we successfully doped ZrNiSn to be p-type and measured the thermoelectric properties up to 850 K. We estimated the band gap size using the Goldsmid–Sharp formula:

\[
E_g = 2eS_{\text{max}}T_{\text{max}}
\]

where \(T_{\text{max}}\) is the temperature at which the maximum of the absolute thermopower \((S_{\text{max}})\) occurs.\textsuperscript{1}

\(E_g\) here is referred to as the thermopower band gap which is obtainable from high temperature thermopower measurements. However, we found a discrepancy between the values obtained from our p-type samples (\(\approx 0.05\) eV) and the n-type results from the literature (\(\approx 0.23\) eV), both of which are different from the \textit{ab initio} calculated result of 0.5 eV.\textsuperscript{25} With the aid of the optical measurements, we used the original derivation of the Goldsmid–Sharp formula to resolve this apparent discrepancy by considering a large difference in weighted mobility between electrons and holes.

### Experimental methods

The solid solution Zr\(_{1-x}\)Sc\(_x\)NiSn \((x = 0, \ldots, 0.10)\) was prepared by arc melting of stoichiometric amounts of Zr (99.99%), Ni (99.99%), Sn (99.999%), and Sc (99.999%) in a water-cooled crucible under an Ar atmosphere. To ensure compositional homogeneity, samples were flipped and remelted five times. The as-cast samples were annealed in evacuated quartz tubes at 1220 K for 7 days, followed by quenching in ice water to ensure the crystalline order. The crystal structure of the samples was studied by X-ray diffraction (XRD) on a Siemens D5000 diffractometer using Cu K\(\alpha\) radiation \((\lambda = 1.5418\) Å). The powder XRD patterns of all samples showed that they were a single phase with cubic C1b structure.\textsuperscript{26}

The electrical resistivity was measured up to 850 K using the Van der Pauw technique. The Hall coefficient, carrier concentration, and mobility were measured in a \(\pm 2\) T magnetic field using pressure contacts.\textsuperscript{27} The Hall carrier concentration \(n_H\) is calculated via:

\[
n_H = -\frac{1}{eR_H}
\]

where \(R_H\) is the Hall coefficient and \(e\) is the electron charge. The Hall mobility, \(\mu_H\), is calculated from the Hall coefficient and electrical conductivity via:

\[
\mu_H = \frac{R_H\sigma}{e}
\]

The thermopower \(S\) was calculated from the slope of \(\Delta V\) \textit{versus} \(\Delta T\) measurements using chromel–niobium thermocouples and by oscillating the temperature gradient \(\pm 10\) K.\textsuperscript{28} The thermal conductivity, \(\kappa\), was measured using a Netzsch laser flash diffusivity instrument (LFA 457) with a Pyroceram® standard. All the samples were coated with a thin layer of graphite to minimize emissivity errors. The data were analyzed using a Cowan model with pulse correction. The heat capacity was estimated using the Dulong–Petit method and the sample densities were calculated from the molar mass and the lattice parameter for each sample obtained from XRD. The optical band gap was measured with the diffuse reflectance infrared Fourier transform spectroscopy method (DRIFTS). A Nicolet 6700 Fourier transform infrared spectrophotometer and a Praying Mantis attachment were used. The detector used was a deuterated triglycine sulfate (DTGS) detector with a KBr beamsplitter. The absorption coefficient was obtained from a Kubelka–Munk analysis.

\[
F(R) = \frac{\alpha}{K} = \frac{(1 - R)^2}{2R}
\]

where \(R\) is the fractional reflectance, \(\alpha\) is the absorption coefficient, and \(K\) is the scattering coefficient. Because we expect that the particle size exceeds the light wavelength, typically measured between 2–20 \(\mu\)m, the scattering coefficient is assumed to be independent of the light frequency.\textsuperscript{29} Burstein–Moss shifts, which result from electron occupation at the band edge, were not considered based on the high thermopower (and therefore the nearly non-degenerate behavior) of the undoped sample.\textsuperscript{30}

### Results and analysis

#### XRD and thermoelectric transport properties

**XRD.** Fig. 1 shows a representative XRD pattern for Zr\(_{0.95}\)Sc\(_{0.05}\)NiSn, which conforms to the well-defined cubic structure typical for the HH phase containing negligible amounts of Sn metal. The calculation of the lattice parameter and the crystal structure refinements were done using the Rietveld method, which yielded a lattice parameter of \(a = 6.1148\) Å for the undoped sample, which agrees with the values reported in the literature.\textsuperscript{29} The inset in Fig. 1 shows a roughly monotonic increase with increasing Sc amount, which
agrees well with Vegard’s law. The increase serves as evidence of the substitution of slightly larger Sc atoms (radius = 1.60 Å) for Zr (radius = 1.55 Å).

Electrical transport properties. The electrical properties of the Zr$_{1-x}$Sc$_x$NiSn ($x = 0, ..., 0.10$) solid solution are shown in Fig. 2a. At room temperature, the samples exhibit high electrical resistivity, which decreases with increasing temperature for all the samples, indicating semiconducting behavior. For high substitution levels, both the room temperature value and the temperature dependence of the resistivity decreases, pointing to a shift of the Fermi level towards the valence band as a result of an increasing p-type carrier concentration.

It can be seen from Fig. 2b that for increasing Sc content in the Zr$_{1-x}$Sc$_x$NiSn solid solution, the Hall mobility ($\sigma R_H$ in units of mobility) is suppressed (and eventually becomes positive) by the addition of holes. The parent compound n-type ZrNiSn exhibits the highest mobility, with a value at room temperature of 25 cm$^2$ V$^{-1}$ s$^{-1}$, which is a typical value for n-type XNiSn-based systems. This value is still significantly lower than 150 cm$^2$ V$^{-1}$ s$^{-1}$ for Bi$_2$Te$_3$ or 900–1400 cm$^2$ V$^{-1}$ s$^{-1}$ for PbTe.

If we assume that each replacement of Zr by Sc leads to one hole as a carrier, the room temperature electron concentration should be completely compensated when the Sc concentration exceeds ~5%. At low Sc content, the samples with 0 ≤ x ≤ 0.04 possess a negative value for $\sigma R_H$, caused by the highly mobile electrons, which is consistent with a substitutional doping explanation.

The carrier mobility depends on the band effective mass $m^*$ and the relaxation time $\tau$ roughly as $\mu = e\tau/m^*$. An estimate of $m^*$ can be obtained from the ab initio calculations of the electronic band structure or from experimental Seebeck and Hall measurements (usually by applying the parabolic band approximation). Since the effective mass varies inversely with the curvature of the bands, a high effective mass is the result of shallow or flat bands, whereas a small effective mass is typical for highly dispersed bands. Based on ab initio calculations, a high mobility value for n-type ZrNiSn is expected owing to their light conduction band (CB). The valence band (VB), however, has been shown to have a higher effective mass, and so by substituting Zr with Sc, one might expect a lower mobilities (in magnitude) for p-type materials as the Fermi energy gets shifted into the VB at $T$. This is consistent with our experimental results, shown in Fig. 2b: the Hall mobilities of the heavily Sc-doped samples are much lower in magnitude than the n-type ZrNiSn undoped sample. In general, the transport properties of materials with small band gaps are influenced by both types of carriers, with a higher weight given for the carriers with higher mobility:

$$R_H = \frac{n_p \mu_p \tau^2 - n_h \mu_h \tau^2}{e(n_p \mu_p - n_h \mu_h)} \tag{5}$$

Here, $n_p$ and $n_h$ are the electron and hole concentrations, respectively, and $\mu_h$ and $\mu_p$ are the electron and hole mobilities, respectively. If $\mu_h \approx \mu_p$, the sign for the Hall coefficient $R_H$ can be negative even if $n_p \geq n_h$. In order to obtain a positive value for the Hall coefficient, $n_p \mu_p$ needs to exceed $n_h \mu_h$. Because the individual carrier contributions, $n_h$ and $\mu_h$, are unknown and difficult to determine, eqn (2) can be used assuming a single carrier type, but it will result in a value that is greater than either the true value for $n_p$ or $n_h$. 

![Fig. 1](image1.png) 

**Fig. 1** Powder XRD pattern for the Zr$_{0.95}$Sc$_{0.05}$NiSn sample (I$_{exp}$) with the Rietveld refinement (I$_{ref}$) and the difference profile (I$_{exp}$−I$_{ref}$). The inset shows the dependence of the lattice parameter $a$ (Å) for Zr$_{1-x}$Sc$_x$NiSn versus Sc content ($x$). The straight line represents the linear fit to Vegard’s law.
The Seebeck coefficient (S) for the Zr$_{1-x}$Sc$_x$NiSn solid solution are presented in Fig. 3. The pure ZrNiSn compound, without any Sc doping, displays a large negative Seebeck coefficient at room temperature, indicating a significant n-type defect concentration in ZrNiSn. With increasing Sc content, the value of the thermopower decreases, as compensating p-type defects (Sc on Zr sites) are added to the naturally n-type material. It can be seen from the change in sign of the Seebeck coefficient from negative to positive that holes become the predominant charge carriers. This sign change occurs at lower Sc contents than observed for the sign change of $\sigma_{\text{RT}}$ (Fig. 2b), indicating that in this region of mixed conduction, the holes probably outnumber electrons. However, because the electrons are more mobile, the Hall coefficient does not change sign until $x > 0.05$ (note that the mobilities are squared, in eqn (5)), whereas the Seebeck coefficient is weighted by mobility to the first power.$^{39}$ As observed in the Hall coefficient measurements, ambipolar conduction of both electrons and holes is most likely responsible for the decrease in the thermopower at high temperatures.

In addition to indicating the onset of bipolar conduction, the maximum of the thermopower can be used for the estimation of the band gap, $E_g$, according to eqn (1),$^{4}$ where we obtained values on the order of 0.05 eV. This is much smaller than the band gap suggested by Aliev et al. from electrical resistivity measurements ($0.18$ eV).$^{46}$ The discrepancy between the estimation from our thermopower data and the literature estimates for the size of $E_g$ in these compounds will be discussed in detail later.

The total thermal conductivity ($\kappa_{\text{tot}}$) in a thermoelectric material is the sum of the electronic contribution ($\kappa_e$) and the lattice contribution ($\kappa_l$). The total thermal conductivity was calculated from the thermal diffusivity ($D$) with $\kappa = C_p D d$, where $C_p$ is assumed to be the Dulong–Petit heat capacity and $d$ the density calculated from the molar mass and the lattice parameter for each sample obtained from XRD. As shown in Fig. 4, the room temperature thermal conductivity is reduced by 40% for Zr$_{0.95}$Sc$_{0.1}$NiSn as compared with the undoped ZrNiSn. Because this is accompanied by a decrease in resistivity (Fig. 2a), it is clear that this must be due to scattering of phonons from disorder induced by Sc substitutions, which reduces the lattice thermal conductivity (which is much greater than the electronic thermal conductivity at room temperature for all Sc compositions). Above room temperature, we can see an increasing thermal conductivity for the doped samples, which is consistent with the existence of both electrons and holes (as evidenced by Fig. 2b and 3). This effect occurs at lower and lower temperatures as the Sc content is increased. Although this is less pronounced for the Seebeck coefficient measurements, thermal conductivity is more sensitive, as it is affected by both the overall decrease of the resistivity and by the bipolar term of the thermal conductivity.

**Optical properties and band gap estimation**

**Band gap measurement.** Optical properties were also measured to obtain information about the band structure, as displayed in Fig. 5. Undoped ZrNiSn was measured in diffuse reflectance at room temperature; the indirect optical band gap was extracted and the Tauc method was applied.$^{36,41}$

$$ (\alpha h\nu)^{1/2} \propto (h\nu - E_g) $$

(6)

![Fig. 5 Optical diffuse reflectance data plotted as the indirect band gap transformation of the Kubelka Munk function for pure ZrNiSn. A linear fit (red dotted line) was used to estimate the band gap by extrapolating to zero absorption, indicating the band gap is ~0.13 eV.](Image)
Here $\alpha$ is the absorption coefficient, $h\omega$ is the photon energy, and $E_g$ is the band gap derived from a linear extrapolation to zero absorption on the plot. The estimated value of the indirect optical band gap was 0.13 eV. Aliev et al. measured similar samples previously, and they reported a minimum in the absorption coefficient of approximately 2000 cm$^{-1}$ (~0.25 eV), a value that is larger than those measured in this work, whereas electrical resistivity measurements gave gaps closer to 0.18 eV.\textsuperscript{48} However, differences in results from those of Aliev et al. and ours could be due to the difference in the measurement technique used. We used DRIFTS, whereas Aliev et al. used transmission/reflection measurements. Nonetheless, we believe that the optical band gap reflects the true gap of the system.

**Estimation of band gap from thermopower measurements.**

When considering the transition from heavily doped to intrinsically semiconducting behavior ($E_{g,th} < \sim 5 \ k_B T$), the Seebeck coefficient is an important indicator of the excitation of minority charge carriers across the band gap. In between these two regions, the thermopower will reach its maximum (as shown in Fig. 3), which can then be used to estimate the band-gap via the Goldsmid–Sharp formula (eqn (1)).

The thermopower band gaps for the Zr$_{1-x}$Sc$_x$NiSn solid solution are compared to those of a series of n-type XNiSn samples from the literature, as shown in Fig. 6.

A large difference in the Goldsmid–Sharp thermopower band gap is apparent when comparing Sc-doped p-type samples from this work and n-type literature results; both are much smaller than those estimated by ab initio calculations (~0.5 eV).\textsuperscript{25,45} Here, we note that even though $X$ (in XNiSn samples) varies through Zr, Hf, and Ti for n-type samples in Fig. 6, the band structures and band gaps are expected to be approximately the same.\textsuperscript{46} All of the n-type samples display a large Goldsmid–Sharp band gap (>0.2 eV) compared to the analogous XNiSn p-type samples. Although the difference might lead one to the conclusion that the choice of dopant will affect the size of the gap, it is important to consider the limitations of the simple Goldsmid–Sharp band gap estimation and the parameters that might affect the results. The Seebeck coefficient for a mixed semiconductor is given as:

$$S = \frac{1}{1 + \frac{\sigma_n}{\sigma_p}} \left( S_p + \frac{\sigma_n}{\sigma_p} S_n \right)$$ (7)

where $S_{n,p}$ and $\sigma_{n,p}$ represent the thermopower and the electrical conductivities of the two different carrier types, respectively. Goldsmid and Sharp derived a simple model assuming classical statistics,\textsuperscript{4} which estimates the electron-to-hole conductivity ratio as

$$\frac{\sigma_n}{\sigma_p} = A \exp(\eta_n - \eta_p)$$ (8)

where $\eta_n$ and $\eta_p$ are the electron and hole dimensionless chemical potentials ($\eta = \xi/k_BT$), respectively, $\xi$ is the chemical potential. $A$ is defined as the weighted mobility ratio.

$$A = \frac{\mu_n N_{n,p}}{\mu_p N_{n,p}} \left( \frac{m^*}{m} \right)^{3/2}$$ (9)

where $\mu_n,p$ is the electron and hole mobility, respectively, $N_{n,p}$ is the valley degeneracy of the electrons and holes, respectively, and $m^*$ is the single valley effective mass (not including the degeneracy of the valence or conduction bands). By substituting eqn (8) into (7), taking the classical limit for the thermopower (eqn (10)), and determining its maximum, Goldsmid and Sharp obtained eqn (11):

$$S = \frac{k_B}{e} \left( \eta_n - r - \frac{5}{2} \right) \left( \eta_n + 2r + 4 \right) - \left( -\eta_n - \eta_p - r - \frac{5}{2} \right)$$ (10)

and

$$A \exp(2\eta_n + \eta_p) = (\eta_n + 2r + 4) \left[ 1 \pm \sqrt{1 - \frac{1}{(\eta_n + 2r + 4)^2}} \right]$$ (11)

where $\eta_g$ is the reduced band gap ($\eta_g = E_{g,th}/k_BT$) and $r$ is the scattering exponent. In light of the large mobility difference between the n-type undoped sample (~25 cm$^2$ V$^{-1}$ s$^{-1}$) and the heavily Sc-doped samples (p-type mobility approximately ~1–2 cm$^2$ V$^{-1}$ s$^{-1}$), a conductivity ratio parameter ($A$) that is larger than unity should be considered. In order to investigate this, we first gathered estimates of the Goldsmid–Sharp band gap for a series of p- and n-type ZrNiSn samples, as shown by the bars in Fig. 6. A clear separation is observed between p-type samples (~0.05 eV) in purple and n-type ones in red (~0.2 eV). For reference, the results from the optical measurements

---

**Fig. 6** Estimation of the band gap for different n-type (red bars) and p-type (purple bars) half-Heusler compounds using the Goldsmid–Sharp formula $E_g = 2eS_{\text{max}} T_{\text{max}}$ in units of eV versus the numerated HH compounds. The dotted line indicates the estimated measured band gap of 0.13 eV for pure ZrNiSn from DRIFTS. The diamonds represent the estimated true band gap value that yields the experimental $S_{\text{max}}$ and $T_{\text{max}}$ when using the full relation derived by Goldsmid and Sharp for an electron-to-hole weighted mobility ratio of $A = 5$. (1) Sc$_{0.03}$Zr$_{0.97}$NiSn, (2) Sc$_{0.04}$Hf$_{0.96}$NiSn, (3) Sc$_{0.02}$Zr$_{0.98}$NiSn, (4) Sc$_{0.02}$Zr$_{0.98}$NiSn, (5) Zr$_{0.02}$Ni$_{0.98}$Sn$_{0.02}$, (6) Zr$_{0.02}$Ni$_{0.98}$Sn$_{0.02}$, (7) Zr$_{0.02}$Hf$_{0.98}$NiSn$_{0.02}$, (8) Zr$_{0.02}$Hf$_{0.98}$NiSn$_{0.02}$, (9) Zr$_{0.02}$Hf$_{0.98}$NiSn$_{0.02}$, (10) Hf$_{1-x}$Ni$_x$Sn$_2$, (11) Ti$_{0.05}$Hf$_{0.95}$NiSn$_{1-x}$, (12) Ti$_{0.05}$Hf$_{0.95}$NiSn$_{0.995}$Sn$_{0.005}$, (13) Ti$_{0.05}$Hf$_{0.95}$Ni$_{0.995}$Sn$_{0.005}$, (14) Ti$_{0.05}$Hf$_{0.95}$Ni$_{0.995}$Sn$_{0.005}$, (15) Hf$_{0.05}$Zr$_{0.25}$Ni$_{0.995}$Sn$_{0.005}$, (16) Hf$_{0.05}$Zr$_{0.25}$Ni$_{0.995}$Sn$_{0.005}$, and (17) Hf$_{0.05}$Ti$_{0.25}$Zr$_{0.75}$NiSn$_{0.995}$Sn$_{0.005}$. The thermopower band gaps for the Zr$_{1-x}$Sc$_x$NiSn solid solution are compared to those of a series of n-type XNiSn samples from the literature, as shown in Fig. 6.
Published on 29 September 2014. Downloaded on 19/02/2015 15:58:44.

Measurements and the band gap estimations. The red line is the result for the n-type doping and the blue line (0.13 eV). The displayed error bars assume an error of the thermopower and the optical band gap. Fig. 7 shows an (0.13 eV). The displayed error bars assume an error of 

Although Goldsmid and Sharp state that even with a large mobility ratio value (\(A = 10\)), the simple band-gap estimation should hold to within \(~\sim 20\%\), we note that they assume a band gap of 10 \(k_B T_{\text{max}}\) to derive this relation. This is a poor assumption for most materials, as the band gap in our system (and many others) is less than 5 \(k_B T_{\text{max}}\). In order to illustrate this point, we have plotted \(2eS_{\text{max}}T_{\text{max}}/E_g\) versus \(S_{\text{max}}\) for a mobility ratio parameter \(A = 5\) using eqn (10) and (11) (Fig. 7). The red line is the result for the n-type doping and the blue line displays the result for p-type doping. Data points in their corresponding colors were taken from their Goldsmid–Sharp band gap ratio to the optical band gap measured in this work (0.13 eV). The displayed error bars assume an error of \pm 10\% in the thermopower and the optical band gap. Fig. 7 shows an increasing \(2eS_{\text{max}}T_{\text{max}}/E_g\) for the n-type samples, whereas it decreases for p-types. Qualitatively, this is because the more mobile electrons dominate the Seebeck coefficient equation because it is weighted by their conductivity; this forces the thermopower of the p-type samples to roll over at lower values, whereas n-type samples maintain higher thermopower at higher temperatures. When looking at recent HH literature including both calculations and experiments—it seems that there is still some uncertainty as to the nature of the valence and conduction bands. Some studies suggest that the smaller observed gaps could be due to impurity doping.\(^\text{57,58}\) Experimental evidence of this is largely based on \(\log(\rho)\) vs. \(1/T\) behavior\(^\text{46,47}\)—the viability of which will certainly be affected by a large electron-to-hole weighted mobility ratio, \(A\). However, because we observe a small optical gap in the undoped ZrNiSn sample, the narrow gap cannot be explained by impurities alone. Several studies suggest that antisite defects or disorder on the Ni sites can lead to impurity bands within the gap.\(^\text{46}\) Recent synchrotron data presented by Xie et al. suggests that partial occupation of the vacant Ni sites occurs in ZrNiSn\(^\text{49,50}\) and the effect is observed in the transport properties. The proposed picture of in-gap Ni states induced by disorder is consistent with the results obtained here. In fact, our discovery of a narrow gap supports this picture since these “in-gap” states, could simply compose the valence band from which we observe optical transitions.

While the inherent very small mobility in p-type ZrNiSn may be disheartening, it beneficially results in suppressed bipolar effects in the n-type ZrNiSn system. Thus, the low mobility minority carrier enables the n-type material to maintain a high thermopower at high temperatures, despite its narrow band gap. This is contrary to the high band gaps found in other p-type HHs, like TiCoSb or the recently identified FeV\(_{0.6}\)Nb\(_{0.4}\)Sb where the band gap estimated from the Goldsmid–Sharp formula is around 0.57 eV and 0.4 eV respectively.\(^\text{51,52}\) Because of the narrow gaps in XNiSn HH materials, one might consider the effect of band non-parabolicity using the two-band Kane model (as applied by Sofo and Mahan).\(^\text{53}\) Although, because DFT calculations show indirect gap behavior—the two-band Kane model should not be applicable.\(^\text{45}\)

Conclusions

In this work, the thermoelectric transport properties of Sc-substituted ZrNiSn HH solid solutions were systematically studied. The substitution of Zr by Sc led to the successful introduction of holes into the system, resulting in a p-type material with a maximum thermopower of +115 \(\mu\text{V} \text{K}^{-1}\) at 650 K. Owing to the introduction of holes into the system, the Seebeck coefficient became positive and increased with increasing temperature, reaching a maximum as the higher mobility n-type carriers were thermally activated. Generally, the transport properties are dominated by the high mobility of the electrons over that of the holes, which can be seen from the \(R_\sigma\) product (in units of mobility). Both p-type data from this work and n-type literature data for the thermopower gap deviated significantly from the optical measurements (0.13 eV), but we have shown that this can be simply explained by a large difference in the weighted mobility between electrons and holes. A high electron-to-hole weighted mobility ratio leads to a suppression of the bipolar effect in the thermoelectric transport properties which is essential for high \(\text{zT}\) values in n-type XNiSn (\(X = \text{Ti}, \text{Zr}, \text{Hf}\)) HH compounds.

Mechanisms to suppress bipolar effects have been discussed to great length in a variety of systems that range from simply

![Fig. 7 A plot of the ratio of the Goldsmid–Sharp band gap (2eS_{\text{max}}T_{\text{max}}) to the true band gap for different p- and n-type half-Heusler compounds in red and blue respectively for an electron-to-hole weighted mobility ratio of \(A = 5\). The solid lines use Goldsmid and Sharp’s full original derivation (eqn (10) and (11)) to estimate this ratio. Experimental points use the observed maximum Seebeck coefficient \(S_{\text{max}}\) and \(T_{\text{max}}\) and the optical band gap \(E_g\text{optical}\). An error of \pm 10\% was assumed for the thermopower measurements and the band gap estimations.](Image 57x149 to 278x296)
changing the band gap or carrier concentration,\textsuperscript{14,15} to preferential minority carrier scattering at grain boundaries or nano-inclusions.\textsuperscript{16,17} This work provides a strategy and method for characterizing bipolar suppression, given a material can be doped both p- and n-type. While there have been efforts to characterize “band engineering” in materials with a single carrier type, the quality factor analysis has not been extended to valence-conduction band systems. In systems such as Bi\textsubscript{2}Te\textsubscript{3}, the most common commercial thermoelectric material, both high and room temperature $zT$ is limited due to minority carrier activation. Strategies to reduce the minority carrier weighted mobility could provide significant impact on the thermoelectric efficiency.

Acknowledgements

We gratefully acknowledge the German BMBF joint project TEG 2020. G.J.S. would like to acknowledge funding from the Bosch-BERN program and The Materials Project: supported by Department of Energy's Basic Energy Sciences program under Grant no. EDCBEE, DOE Contract DE-AC02-05CH11231 for their funding. The authors would also like to acknowledge the Molecular Materials Research Center (MMRC) at the California Institute of Technology for use of their optical instruments for measurements in this work. Moreover, we would like to thank Chirranjeevi B. Gopal, Julia Krez, and Fivos Drymiotis for helpful discussions.

Notes and references